

Probabilistic forecasting of geomagnetic indices using solar wind air mass analysis

Robert L. McPherron

Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California, USA

George Siscoe

Center for Space Physics, Boston University, Boston, Massachusetts, USA

Received 10 June 2003; accepted 15 September 2003; published 9 January 2004.

[1] Measurements of the Sun's photospheric magnetic field can in principle be used to predict geomagnetic activity 1–3 days in advance. The accuracy of such predictions is low, however, because they do not include the north-south component of the interplanetary magnetic field (IMF B_z), which is the most geomagnetically relevant parameter. Aside from that carried by large-scale transients, IMF B_z is mainly a product of small-scale, in-transit turbulence and so is inherently unpredictable from solar measurements. Routine 1- to-3-day forecasts of geomagnetic activity based on deterministic algorithms are, therefore, not possible. Probabilistic forecasts offer the next best thing to deterministic forecasts, and air mass climatology offers a way to develop the advantages inherent in probabilistic forecasts for space weather applications. Here we address the IMF B_z indeterminacy problem (or better, get around it) by applying the concept of air mass climatology to the solar wind. We give criteria and statistics for solar wind air masses, which provide proof of concept for routine, midrange (1- to 3-day) probabilistic air mass forecasts of daily levels of geomagnetic activity. *INDEX TERMS*: 2722 Magnetospheric Physics: Forecasting; 2164 Interplanetary Physics: Solar wind plasma; 2134 Interplanetary Physics: Interplanetary magnetic fields; 2788 Magnetospheric Physics: Storms and substorms; *KEYWORDS*: space weather, forecasting, geomagnetic indices

Citation: McPherron, R. L., and G. Siscoe (2004), Probabilistic forecasting of geomagnetic indices using solar wind air mass analysis, *Space Weather*, 2, S01001, doi:10.1029/2003SW000003.

1. Introduction

[2] This paper addresses the problem of extending routine forecasts of geomagnetic disturbance indices beyond the order of 1 hour allowed by solar wind measurements taken at the forward Langrangian point (L1). Forecasts based on solar measurements, such as the Wang-Sheeley model [Wang and Sheeley, 1997], can increase forecast time up to the 3 or 4 days that it takes the solar wind to travel from the Sun to the Earth. Longer-range forecasts rely on 27-day recurrence of solar conditions that are correlated with geomagnetic activity. Someday, solar physicists might be able to predict from present solar measurements future solar conditions that are correlated with geomagnetic activity; this capability will then also extend forecasting range beyond the solar wind transit time. To have a way to distinguish the three ranges of forecast potential, L1-to-Earth, Sun-to-Earth, and beyond-Sun-to-Earth, we will refer to them respectively as short-range, midrange, and long-range forecasts.

[3] We report here first results of a new technique to maximize the accuracy of routine, midrange forecasts of geomagnetic disturbance indices based on solar measurements. We illustrate the technique by providing graphs that could be used now with the Wang-Sheeley model as modified by Arge *et al.* [2003] to forecast the daily magnetic disturbance index (A_p). (We will refer to Arge-modified

Wang-Sheeley model as the Wang-Sheely-Arge (WSA) model.) The new technique addresses the following basic problem of midrange forecasting. The southward component (B_s) of the interplanetary magnetic field (IMF) or, more accurately, the eastward component (E_e) of the interplanetary electric field (IEF with $E_e = VB_s$, where V is solar wind speed) is the main controller of geomagnetic disturbance [e.g., Reiff and Luhmann, 1986]. Except for interplanetary coronal mass ejections (ICMEs) and seasonal dipping of the Parker spiral magnetic field into the Northern or Southern Hemisphere of the tilted geomagnetic dipole, which gives rise to the Russell-McPherron effect [Russell and McPherron, 1973], this controlling quantity is virtually impossible to predict from solar measurements. The reason for this inherent unpredictability is easy to state. The IMF alternates irregularly but on average <10 min between geomagnetic north and south. Outside of ICMEs, which can maintain their north-south magnetic structure [e.g., Zhao and Hoeksema, 1997, 1998], at any time there are typically more than 700 north-south alternations of the IMF (and so east-west alternations of E_e) between the Sun and the Earth. Imagine predicting on the basis of solar measurements the structure of any aspect of the solar wind between the Sun and the Earth with a resolution the order of 0.1%.

[4] Things that are predictable from solar measurements (solar wind speed, the Parker spiral, sector polarity, IMF

strength, and (in principle, at least) ICME magnetic structure) have much bigger spatial scales. These quantities are governed by large-scale dynamics, which means that they change deterministically, if at all, between the Sun and the Earth. By contrast, the north-south component of the IMF appears to be governed by the dynamics of small-scale turbulence and so is inherently unpredictable. This is the problem that confronts anyone who wants to use the north-south component of the interplanetary magnetic field (IMF B_z) to do midrange forecasting. We will call it the B_z indeterminacy problem, where, as conventional, z represents the north-south component of the IMF. ($B_z = -B_s$ in the GSM coordinate system, but B_z can also be defined in the GSE coordinate system, which is used to generate statistics for this study.)

[5] Obviously, the B_z indeterminacy problem must impact a forecaster's ability to make routine midrange predictions of geomagnetic activity; the daily A_p index, for example. Because of it, forecasters base tomorrow's and subsequent days' predictions of A_p not on IMF B_z (except for the Russell-McPherron effect) but instead on other, mostly indirect indicators of geomagnetic activity that are not directly tied to IMF B_z . The new technique described here gives a method for dealing with the IMF B_z indeterminacy problem (or getting around it) in making routine, midrange predictions of geomagnetic activity. Of course, there is no way to avoid the consequences of the inherent unpredictability of IMF B_z from solar measurements. The consequence here is that the forecaster must make probabilistic rather than deterministic predictions. The forecast variable that the proposed technique produces is a plot of the predicted 3-hour a_p index as a function of probability. Below we give examples in the form of plots of the probability that a_p will exceed a specified value. The forecaster can then easily prepare as a forecasting product the median daily A_p or its modal value, in each case giving as a measure of the uncertainty in the prediction the standard deviation or the quartiles. He or she can also give the probability that the index will exceed some threshold, for example, "there is a 25% chance that $A_p > 50$ nT."

[6] The project is to use archival data to derive statistics of solar wind speed, IMF B_z , and IEF y component of the IEF ($E_y = -VB_z$) tied to specific solar wind conditions that are midrange predictable from solar measurements. In meteorology, such a procedure would correspond to compiling air mass climatologies; for example, maritime-tropical air is warm and humid, continental-polar air is cool and dry, etc. The procedure is useful to the extent that the air masses are well separated in their meteorologically relevant qualities and to the extent that each air mass has well-defined statistical properties. In the atmosphere an air mass is defined by its origin. The Gulf of Mexico generates maritime-tropical air masses, central Canada generates continental-polar air masses, etc. The boundaries between air masses constitute weather fronts, where storms arise. A map showing the location of air masses and the fronts between them allows a weather forecaster to

predict that locations well inside an air mass will experience air-mass-type weather and that locations near their boundaries will experience frontal-type weather. Since the air mass concept has served traditional meteorology well, it might also be helpful in space weather meteorology.

[7] The following questions bear on whether the air mass concept might be advantageous to space weather meteorology. First, are there bounded volumes of solar wind with uniform statistical geoeffective properties that would correspond to a solar wind analog of an air mass? Second, if so, is there more than one such type of volume, each having distinct statistical properties such that the concept of identifiable air mass types would make sense? Third, if so, can the presence of these air mass types at 1 AU be predicted from solar measurements? If the answer to all three questions is "yes," then the concept of air mass climatology might be useful in space weather forecasting. To this end, one needs to develop the relevant climatologies pertaining to each type of solar wind air mass.

[8] We do not need to invent candidates for air mass types in the solar wind inasmuch as the solar wind is already described in terms of distinct regimes, chief among which are corotating fast streams, corotating slow streams, and ICMEs. (For brevity in the following, we omit "corotating" when referring to fast and slow streams.) Besides these common regimes we have the interface between a leading slow stream and a trailing fast stream, which is the much-studied corotating interaction region (CIR, the subject of an edited monograph [Balogh *et al.*, 1999]). For now, however, we propose to refer to a CIR as a "corotating compression ridge" (CCR), so that we may consistently also refer to its opposite, a "corotating rarefaction slope" (CRS), this being the interface between a leading fast stream and a trailing slow stream. A CCR corresponds to a compressional weather front between fast stream and slow stream air masses [Siscoe, 1972]. Figure 1 gives the ideal geometry relating to fast and slow streams and the compression and rarefaction zones between them [from Pizzo, 1978]. The words "ridge" and "slope" refer to the relative sizes and profiles of the pressure and speed of CCRs and CRSs respectively. Both CCRs and CRSs might be expected to generate distinct types of frontal space-weather conditions.

[9] The present study is in nature similar to a project that B. T. Tsurutani and W. D. Gonzalez [e.g., Tsurutani and Gonzalez, 1997] have pursued over many years to identify the interplanetary causes of magnetic storms. A basic difference, however, is that here the forecast objective is not magnetic storms but rather routine, midrange predictions of daily levels of geomagnetic activity.

[10] Since this is the first exploration of the usefulness of the air mass project to space weather, we have chosen to study as uncomplicated a situation as possible, namely, corotating fast streams, slow streams, and the compression zone between them (CCR). The goal is to determine whether these features satisfy criteria stated earlier for being usefully regarded as solar wind air masses and associated fronts. That is, do they designate solar wind

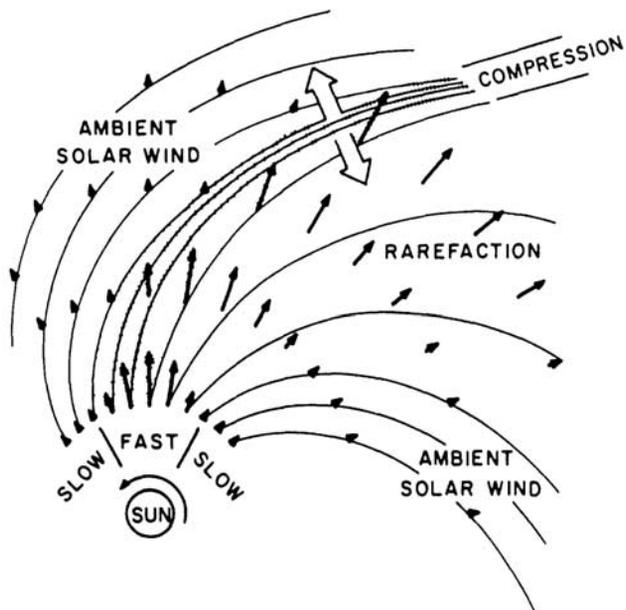


Figure 1. Sketch to illustrate geometry of corotating fast and slow solar wind streams and the corotating compression ridge between them. Thick arrows in outline show the direction of the pressure force that builds up along the corotating compression ridge [from Pizzo, 1978].

conditions at 1 AU that can be predicted from solar measurements and that have distinct, geomagnetically relevant statistical properties? We do not address here the corresponding question for the ICME air mass and its weather front, which is called the “ICME magnetosheath.” In the spirit of keeping the study uncomplicated we restrict the present study to data from the year 1995, when fast and slow streams were well developed and long lived, and the necessary solar, interplanetary, and geomagnetic data were readily available. In this year the situation was nearly ideal for studying air mass statistics of corotating fast and slow streams. Complications resulting from transients were at a minimum.

2. Stream Interface as Marker

[11] We first tried to differentiate solar wind air masses by speed; that is, we assumed that a given speed would identify whether one is in a fast stream or a slow stream. If this were so, then a speed prediction from the WSA model, for example, would allow a forecaster to immediately make a prediction based on the appropriate fast stream or slow stream statistics. This idea, however, failed. A histogram of speed for 1995 does indeed show two peaks, one near 350 km/s and one near 650 km/s (see Figure 2). However, the valley between the peaks is shallow; moreover, in other years, already in 1996, there is only one peak. This means that speed by itself would give many false identifications of air mass type.

[12] We decided that the safest way to know whether one is in a fast stream or a slow stream is to locate the

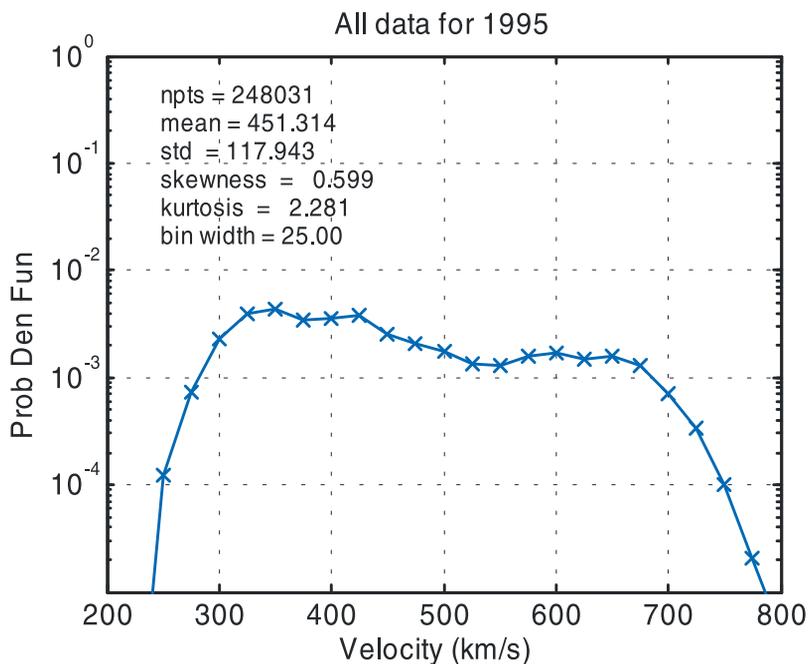


Figure 2. Plot of the probability density function (i.e., the probability that the speed lies between solar wind speed (V) V km/s and $V + 1$ km/s) for solar wind speed for 1995.

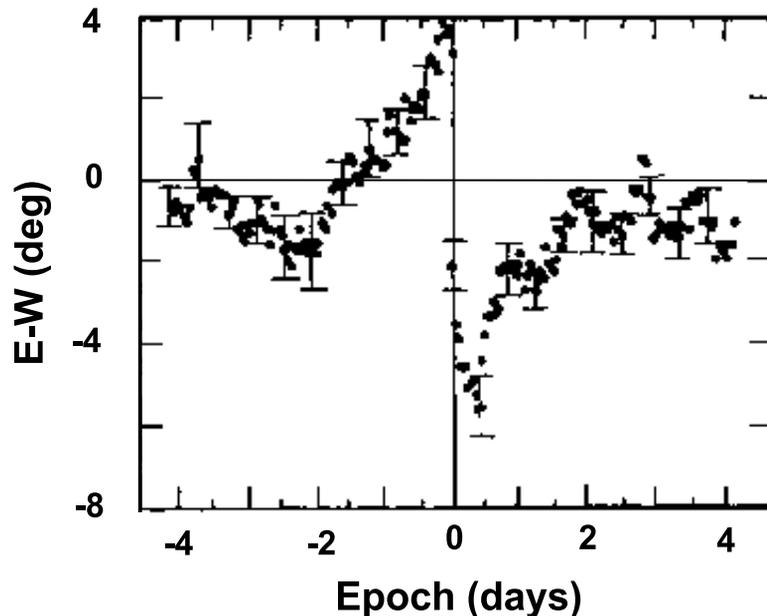


Figure 3. A superposed epoch average of the deflection of solar wind flow in the east-west direction (in degrees) for the year 1994 showing the bipolar signature characteristic of the interaction between corotating slow (leading) and fast (trailing) streams [from *Gosling et al., 1978*].

boundary between the two types. Then prior to the arrival of the boundary, one is assuredly in the leading type of air mass, and after the boundary passes, one is assuredly in the trailing type of air mass. This is true even if the speed of the leading air mass falls in the range of the speed associated with the trailing air mass or vice versa. For the intended purpose, the interface within the CCR, which already has a name by tradition, the “stream interface” [Burlaga, 1974], works better than the interface within the CRS, because it is much better defined in measurements at 1 AU. As importantly, a forecaster can use the WSA model to predict the arrival time of the stream interface from solar data. Accordingly, we used the stream interface as the marker to identify preceding slow stream and following fast stream air mass types.

[13] The issue of the accuracy with which a forecaster can predict the arrival time of the stream interface from the WSA model is important, of course. As a rule, the stream interface lies at the foot of the rise in solar wind speed that leads to the fast stream. So even without refinement, one can predict its arrival about as well as one can predict a change in solar wind speed. The uncertainty in predicting a change in solar wind speed is typically 8 hours, roughly, although the proper statistical study to determine this number still needs to be done. The issue of determining the accuracy with which the arrival time of the stream interface can be predicted will be addressed in a separate study which will specify a prediction procedure that uses the WAS model and also takes account of precursor signals. This study will also provide statistics on the occurrence frequency of false positives

(predicted interfaces that never arrive) and false negatives (stream interfaces that arrive unpredicted).

3. Superposed Epoch Analyses

[14] To begin our test of the usefulness of the air mass concept in the space weather context, we carried out superposed epoch analyses of geomagnetically pertinent quantities using the stream interface as the common origin for the epochs. We identified stream interfaces by a characteristic bipolar east-west deflection of the solar wind flow that inevitably arises when a fast stream pushes against a preceding slow stream along a Parker-spiral interface [Dessler, 1967; Carovillano and Siscoe, 1969]. This bipolar deflection is indicated schematically in Figure 1. Figure 3 shows it as revealed in a classical superposed epoch analysis of 23 interfaces selected for discontinuous density and temperature signatures. These occurred in streams from 1971 to 1974, which include the years of “giant streams” [Gosling et al., 1978]. In our case, we identified 26 such deflections from the 1995 measurement interval. Although all bipolar deflections in 1995 were well defined, the sizes of the streams themselves as gauged by the change in flow speed varied considerably. Unlike 1974, 1995 was not a year of giant streams.

[15] Figure 4 shows the results obtained for five superposed epoch analyses of the 26 identified streams of 1995. From top to bottom, the five quantities in the panels are east-west flow deflection, solar wind speed (V), IMF B_z (in the GSE coordinate system, not the GSM system that is relevant to the Russell-McPherron effect, which is a different topic than the one on statistical properties being

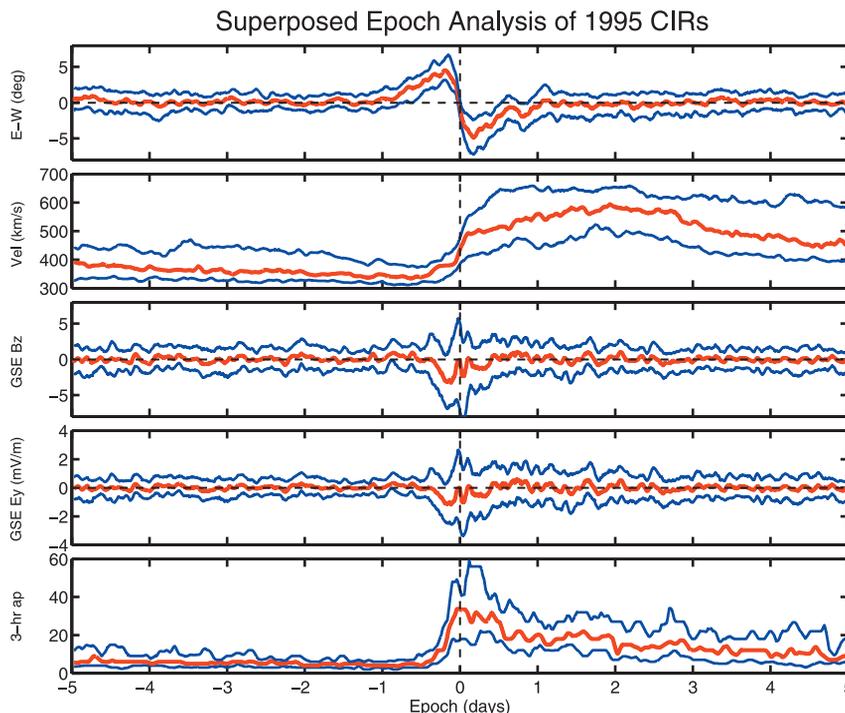


Figure 4. Superposed epoch averages for 26 streams in 1995 of the following quantities (top to bottom): east-west flow deflection; flow speed, north-south component of the interplanetary magnetic field (IMF B_z) (in the GSE coordinate system); the east-west component of the motional electric field ($E_y = -VB_z$); and the 3-hour ap index of geomagnetic activity (ap). Zero on the epoch axis corresponds to the stream interface as identified by the zero crossing of the bipolar east-west flow deflection.

treated here), the geoeffective component of the IEF field ($E_y = -VB_z$), and the 3-hour ap index. Each panel gives the mean and the quartiles for the plotted parameter.

[16] Comparing the top panel with Figure 3 suggests that, on average, the zone of east-west deflections in 1974 might have been twice as wide as in 1995 (4 days versus 2 days). Perhaps giant streams have wider interaction fronts than normal streams, or, alternatively, the real interaction front in the 1974 streams was confined to the rapid change in deflection that occurred during 1-day intervals before and after the stream interface. Either way, the average deflection is $\sim 4^\circ$ in both the 1974 and the 1995 cases.

[17] Several important lessons relative to our purpose can be drawn from these plots. Consider first the difference in the properties of the “pure” fast and slow streams, i.e., their noninteracting parts. These lie ~ 1 day outside of the CCR on both sides of the stream interface as identified by the east-west deflection panel. Comparing the median solar wind speed 1 day before and after the CCR shows that the average fast stream in 1995 was about twice as fast as the average slow stream (670 km/s versus 340 km/s). An interesting difference between the statistical properties of fast and slow streams is the relatively greater variability in the speed of fast streams, evidenced by a considerably greater separation between the quartiles for fast streams.

The quartiles are about a factor of 2 farther apart on the day after than on the day before the CCR.

[18] The median value of IMF B_z 1 day before and after the CCR is ~ 0 , as expected. Presumably, therefore, the sample size (26 streams) is big enough for the expected cancellation of positive and negative values to be essentially complete. Unexpected, however, even surprising, is the absence of a difference between fast and slow streams in the amplitude of the variability of IMF B_z . Outside of the CCR the separation of the IMF B_z quartiles is about the same for fast streams as for slow streams.

[19] The behavior of solar wind speed and IMF B_z contribute jointly to the behavior of E_y , which, as mentioned, is the parameter most directly responsible for geomagnetic activity. The fourth panel shows that, like IMF B_z , the median value of E_y is virtually 0 in both fast and slow streams. Unlike IMF B_z , however, E_y is considerably more variable in fast streams than in slow streams. Its greater variability in fast streams exposes the magnetosphere more often to intervals of intense geoeffective electric field, thus accounting in part for the greater average level of geomagnetic activity during fast streams seen at the bottom of Figure 4. A direct comparison between panels 4 and 5 in Figure 4 is not strictly valid, however, inasmuch as panel 4 gives E_y in the GSE system

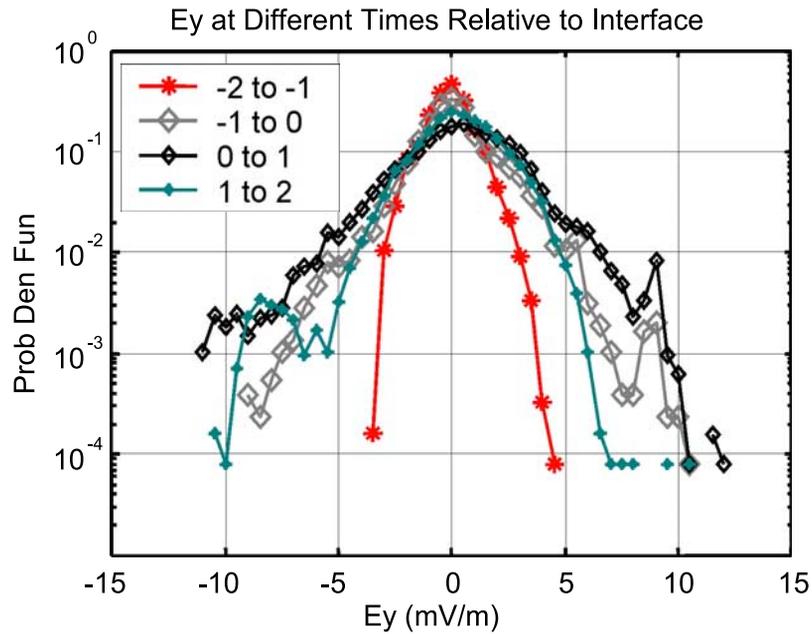


Figure 5. Plot of the probability density function of E_y (i.e., the probability that the field lies between E_y mV/M and $E_y + 1$ mV/m) for four different air mass situations defined by their timing relative to the stream interface.

whereas a_p responds to E_e , which is E_y in the GSM system. The difference can be big because of the Russell-McPherron effect, which is a product of E_e and not E_y , strongly modulates the sustained activity associated with high-speed streams [Crooker and Siscoe, 1986; Crooker and Cliver, 1994]. Panels 2, 3, and 4 of Figure 4 show that were it not for the Russell-McPherron effect, the greater geoeffectiveness of fast streams relative to slow streams would have more to do with the trivial fact that fast streams are faster than slow streams than with fast streams having stronger average southward IMFs. This result might bear on the finding of Crooker *et al.* [1977] relating to the solar wind in general (that is, not separated into fast and slow streams) that A_p correlates best with solar wind speed. IMF B_z plays an important but lesser role [Crooker and Gringauz, 1993].

4. Probabilistic Forecast Relative to Stream Interface

[20] Now that we have a marker predictable from solar measurements (the stream interface) with respect to which the geoeffective element of the solar wind (E_y) seems to be organized, the task is to quantify the statistics of this element relative to the marker. Figure 5 shows the probability density function (PDF) of E_y (i.e., the probability that E_y lies between E_y millivolts per meter (mV/m) and $(E_y + 1)$ mV/m) for 1-day intervals starting 2 days before the stream interface. The interval -2 to -1 lies in the “pure” slow stream, the interval 1 to 2 in the “pure” fast stream, and the two intervals between -1 and 1 in the CCR.

[21] Positive values of E_y correspond to geomagnetically disturbed intervals. We see that the slow stream PDF

shows the lowest probability of significant disturbance. For $E_y > 3$ mV/m, the difference in the PDF for the fast stream exceeds by more than an order of magnitude that for the slow stream. For $E_y > 4$ or 5 mV/m, the PDFs for the CCR exceed that for the fast stream. This means that strong disturbance values of E_y are most likely to occur in the CCR, which accords with findings [Hirshberg and Colburn, 1973; Garrett *et al.*, 1974] that were long ignored after the discovery of coronal holes but which recently have been revived [Crooker and Cliver, 1994; Tsurutani *et al.*, 1995]. The geoeffectiveness of CCRs is evident at the bottom of Figure 4.

[22] The big differences in PDFs that Figure 5 reveals is responsible for correspondingly big differences in each interval’s individual tendency toward geomagnetic calm or disturbance. As an example, Figure 6 contrasts these tendencies as expressed in the cumulative probability distribution function for the a_p index, which gives the probability that a_p will exceed the value on the abscissa. The “curves” (the steps in the curves result from a_p having only 27, discrete values) are based on data recorded during slow streams (1–3 days before the stream interface) and CCRs (2 days centered on the stream interface). For each situation (slow stream or CCR), three curves are given, one corresponding to all data and the other two corresponding to cases in which the magnetic field was weaker than or stronger than the median field strength for all cases. These curves are designated as “low B ” and “high B ,” respectively. The three curves for each situation relative to the stream interface are well separated. The probability that a_p will exceed a given value (bigger than about 20 nT) is an order of magnitude or more greater in CCRs than in slow

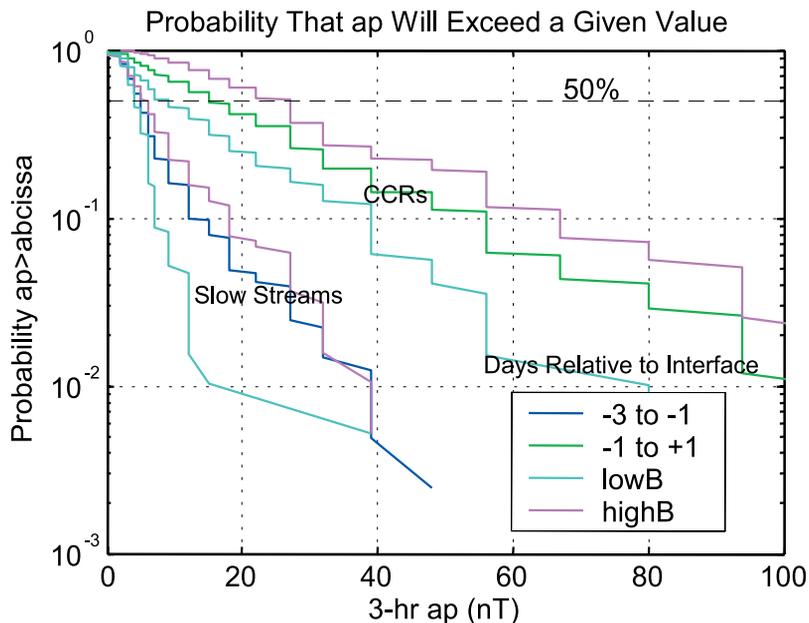


Figure 6. Probabilistic curves for ap for slow streams and for corotating compression ridges. Each case shows three curves corresponding to all data and high and low B field strengths.

streams. Within each situation (slow stream or CCR), high B conditions are more geoeffective than low B conditions, as one would expect.

[23] Figure 7 adds information on the variability of geoeffectiveness within each situation (slow stream or CCR) resulting from variability of solar wind speed. Again, there are three curves for each situation, one corresponding to all data (the same curves as in Figure 6) and the other two

corresponding to cases in which the solar wind was slower than or faster than the median speed for all cases. These curves are designated as "low V" and "high V," respectively. Here, too, one sees the expected result that, within each situation, faster streams are more geoeffective than slower streams. Obviously, one can subdivide the data further into low V, low B; low V, high B; high V, low B; and high V, high B; and one should so divide the data to build

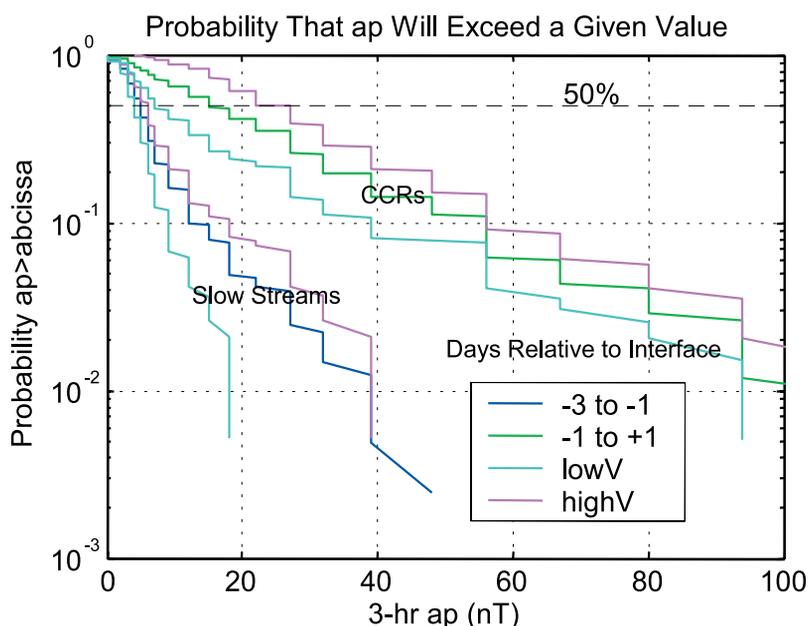


Figure 7. Same as Figure 6 except that the three curves for each case correspond to all data (as in Figure 6) and high and low flow speeds.

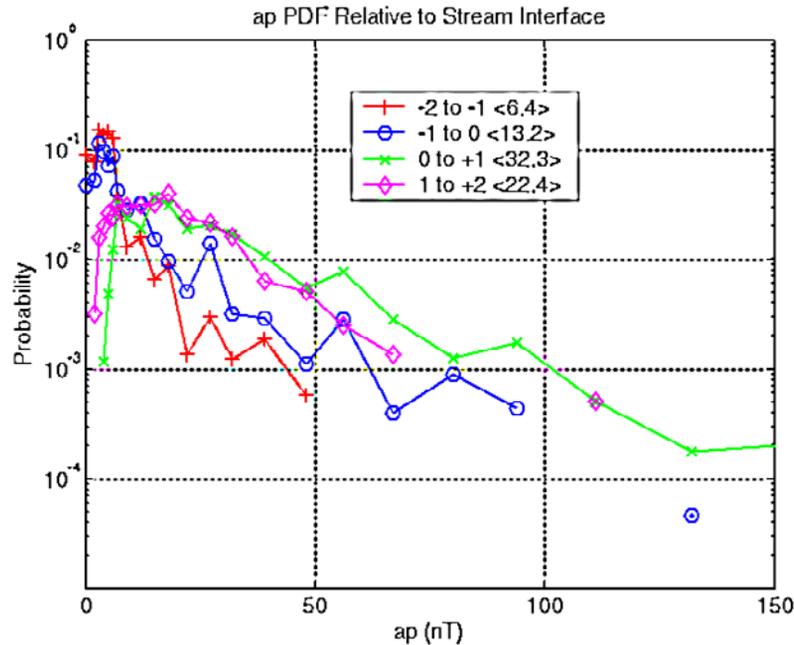


Figure 8. Probability distribution functions (normalized histograms) for ap for four 1-day intervals starting 2 days prior to the stream interface.

statistics of greater forecast accuracy. Plots similar to those in Figures 6 and 7 in which the fast stream replaces the slow stream show similar tendencies except fast stream curves are not as well separated from the CCR curves, which is consistent with the result shown in Figure 5.

5. Payoff

[24] In situations with well-defined fast streams, slow streams, and CCRs, a set of individual plots like those in Figures 6 and 7 for the four combinations of high and low V and B would constitute a new and powerful forecast tool to use to predict tomorrow's or the next day's daily Ap value from the WSA model. Consider by way of illustration the horizontal, dashed line near the top of Figures 6 and 7. It is the 50% probability marker, which means that there is a >50% chance that ap will exceed values on the curves above the line and a <50% chance that ap will exceed values on the curves below the line. To illustrate, it shows that there is a 50% probability that ap will not exceed 4 nT in a low B slow stream, whereas there is a 50% probability that ap will exceed 27 nT in a high B CCR. More lines can be drawn at operationally significant probability values that are appropriate to specific customers (power companies, for example).

[25] Other displays, still based on position relative to the stream interface, might be designed to "package" forecast-enabling information in useful forms. Besides cumulative probability distribution functions like Figures 6 and 7, one can plot the PDFs for the ap index for each air mass situation, as in Figure 8. The PDF is the same as a normalized histogram. Hence from such plots one can

predict the mean, median, and mode of ap , and some measure such as standard deviation or quartiles of the prediction's uncertainty.

[26] Figures 6, 7, and 8 show that once we have determined how to identify a solar wind air mass, we no longer need to worry, from an operational perspective, about solar wind parameters. We can go straight from a prediction of an air mass type and its quantification to probability curves for tomorrow's and the next day's ap index. This is the operational, space weather payoff. The air mass prediction algorithms can be expanded in scope and improved in accuracy as archival and accumulated data allow the statistical characterization of each air mass condition to be refined.

[27] There is, however, a payoff that can be had from air mass statistics of IMF Bz or IEF Ey . While it is not possible to predict the actual waveform of IMF Bz or IEF Ey from solar measurements, it is possible to predict their statistical properties, as this work has shown. From these it is in principle possible to generate surrogate waveforms with the same statistical properties (mean, mode, standard deviation, autocorrelation function, etc.). A surrogate waveform for a given type of air mass situation can be used as input to a numerical, magnetospheric space weather code to test its skill at producing magnetic activity statistics corresponding to the actual statistics for that situation.

6. Summary

[28] Forecasts based on air mass climatology had served traditional meteorology well for decades until replaced by powerful numerical models of atmospheric circulation,

which are now the weather forecaster's main prediction tool. Space meteorology is still in its prenumerical model phase. It is in what, by analogy with the development of traditional meteorology, ought to be its air mass climatology phase.

[29] Air mass climatology as it applies to space weather is what this paper has addressed. Its use in the space weather arena first recommended itself to the authors as a stratagem to advance against what otherwise appears to be an insurmountable upper limit to the accuracy of routine, midrange forecasts of geomagnetic activity. This upper limit leaves much room for improvement, if improvement were possible. The barrier to improvement has been the inability to predict IMF B_z at 1 AU from solar measurements. Were such predictions possible, routine, midrange forecasts could be based on deterministic forecast algorithms and, therefore, would be susceptible to improvement. There is at present no way known even in principle, however, to meet this requirement, since on timescales relevant to midrange forecasts, IMF B_z appears to be predominantly a product of in-transit turbulence having a characteristic autocorrelation time of <1 hour (ICMEs and the Russell-McPherron effect excepted). In response to this predicament, routine probabilistic midrange forecasts based on air mass climatology offer the next best thing to deterministic forecasts. The idea is to identify solar wind air masses that have the following characteristics: They should have stable statistical properties over midrange timescales, they should have well-differentiated statistical properties, and they should be predictable at 1 AU from solar measurements, e.g., by the WSA model.

[30] Candidate solar wind air mass types are fast streams, slow streams, and interplanetary mass ejections (ICME). In considering space weather predictable from these air masses, however, one must also include the interactions between them, analogous to atmospheric weather fronts. These create distinct interaction zones with their own statistical properties. Associated with these three air mass types there are three kinds of interaction zones: a corotating compression ridge (CCR) between a leading slow stream and a following fast stream, a corotating rarefaction slope (CRS) between a leading fast stream and a following slow stream, and an ICME magnetosheath that precedes an oncoming ICME. There might also be a situation in the wake of an ICME that satisfies the criteria for an air mass.

[31] In the spirit of a pilot study, this paper has developed the air mass concept as it pertains to fast streams, slow streams, and CCRs, which are relatively numerous and easy to identify. To approximate ideal conditions, the study has used data from 1995, a year characterized by well-developed, long-lived solar wind streams, relatively few ICMEs, and by the availability of requisite solar, solar wind, and geomagnetic data. A breakthrough that the study has made in applying the air mass concept to fast streams, slow streams, and CCRs is that the stream interface is key to making the concept work. That is, from solar measurements (WSA model), the best predictor of the stated three conditions is position relative to the stream

interface. Using this stream-interface-based method for identifying air mass conditions, the study has shown that the three criteria for the air mass concept to be of use in space weather, stable statistics, distinct statistics, and predictability from solar measurements, are satisfied by the fast and slow stream types and their compressional interface. The study has provided examples of graphs that could be used now as working tools by forecasters to predict the probability of ap values 1–3 days in advance using the WSA model to give values of solar wind speed and magnetic field strength needed as input to the graphs and to predict the arrival times of stream interfaces.

7. Next Steps

[32] In the project to develop the predictive advantages to be had by applying the air mass concept to space weather, this paper has provided proof of concept. Steps beyond proof of concept include the following: (1) Complete the set of graphs illustrated in Figures 6 and 7 (i.e., compile separate graphs for each of the 3 days prior to and following the stream interface for each of the four combinations of high and low speed and field strength and for the two "seasons," equinox and solstice, and for the equinox case, one set of graphs for active R-M effect and one set for inactive R-M effect). (2) Carry out a statistical study of the reliability (truth table) and accuracy of determining the arrival time of stream interfaces from solar data (this study will incorporate stream-interface precursors, such as a drop in alpha-particle flux ahead of the CCR [compare *Gosling et al.*, 1981], to update the forecast using L1 data). (3) Extend the study to other phases of the solar cycle and monitor the change in its reliability and accuracy. (4) Carry out a study with the same goals as step 1 for the ICME magnetosheath. (The ICME itself appears to be amenable to forecast algorithms based on deterministic predictions.) (5) Develop technique to generate surrogate waveforms with the same statistical properties as determined for actual air mass situations.

[33] In step 3 we should anticipate that the number of opportunities per year in which the technique described here might be applied will vary with the solar cycle. Most corotating stream interfaces occur during the declining phase of solar activity, such as 1974 and 1995, when coronal holes, the progenitors of stream interfaces, extend to low latitudes and transient events, which disrupt stream interfaces, are relatively few. This is an optimal situation for applying forecasting techniques tied to stream interfaces. The situation is less favorable for such techniques when solar activity is at its maximum. Then coronal holes and the streams they generate are small, irregular, and short-lived, and stream-disrupting transient events are relatively numerous.

[34] **Acknowledgments.** We thank Nancy Crooker for helpful discussions and useful information. This work was supported in part by the National Science Foundation under its Science and Technology Centers program, grant OIA-0120950 supporting the Center for Inte-

grated Space Weather Modeling (CISM). RLM would like to thank the Boston University Center for Space Physics for its support while on sabbatical leave at Boston University. He also acknowledges support by the National Science Foundation under a Space Weather Grant NSP ATM 02-08501.

References

- Arge, C. N., D. D. Odstrcil, V. J. Pizzo, and L. R. Mayer (2003), Improved method for specifying solar wind speed near the Sun, *AIP Conf. Proc.*, in press.
- Balogh, A., J. T. Gosling, J. R. Jokipii, and H. Kunow (1999), *Corotating Interaction Regions, ISSI Space Sci. Ser.*, vol. 7, Kluwer Acad., Norwell, Mass.
- Burlaga, L. F. (1974), Interplanetary stream interfaces, *J. Geophys. Res.*, 79, 3717–3725.
- Carovillano, R. L., and G. L. Siscoe (1969), Corotating structure in the solar wind, *Sol. Phys.*, 8, 401.
- Crooker, N. U., and E. W. Cliver (1994), Postmodern view of M-regions, *J. Geophys. Res.*, 99, 23,383–23,390.
- Crooker, N. U., and K. I. Gringauz (1993), On the low correlation between long-term averages of solar wind speed and geomagnetic activity after 1976, *J. Geophys. Res.*, 98, 59–62.
- Crooker, N. U., and G. L. Siscoe (1986), The effect of the solar wind on the terrestrial environment, in *Physics of the Sun*, edited by P. A. Sturrock, pp. 193–249, D. Reidel, Norwell, Mass.
- Crooker, N. U., J. Feynman, and J. T. Gosling (1977), On the high correlation between long-term averages of solar wind speed and geomagnetic activity, *J. Geophys. Res.*, 82, 1933–1937.
- Dessler, A. J. (1967), Solar wind and interplanetary magnetic field, *Rev. Geophys.*, 5, 1–41.
- Garrett, H. B., A. J. Dessler, and T. W. Hill (1974), Influence of solar wind variability on geomagnetic activity, *J. Geophys. Res.*, 79, 4603–4610.
- Gosling, J. T., J. R. Asbridge, S. J. Bame, and W. C. Feldman (1978), Solar wind stream interfaces, *J. Geophys. Res.*, 83, 1401–1412.
- Gosling, J. T., G. Borrini, J. R. Asbridge, S. J. Bame, W. C. Feldman, and R. T. Hansen (1981), Coronal streamers in the solar wind at 1 AU, *J. Geophys. Res.*, 86, 5438–5448.
- Hirshberg, J., and D. S. Colburn (1973), Geomagnetic activity at sector boundaries, *J. Geophys. Res.*, 78, 3952–3957.
- Pizzo, V. (1978), A three-dimensional model of corotating streams in the solar wind: 1. Theoretical foundations, *J. Geophys. Res.*, 83, 5563–5572.
- Reiff, P. H., and J. G. Luhmann (1986), Solar wind control of the polar-cap voltage, in *Solar Wind-Magnetosphere Coupling*, edited by Y. Kamide and J. A. Slavin, pp. 453–476, Terra Sci., Tokyo.
- Russell, C. T., and R. L. McPherron (1973), Semiannual variation of geomagnetic activity, *J. Geophys. Res.*, 78, 92–108.
- Siscoe, G. L. (1972), Structure and orientation of solar wind interaction fronts: Pioneer 6, *J. Geophys. Res.*, 77, 27.
- Tsurutani, B. T., and W. D. Gonzalez (1997), The interplanetary causes of magnetic storms: A review, in *Magnetic Storms, Geophys. Monogr. Ser.*, vol. 98, edited by B. T. Tsurutani et al., pp. 77–89, AGU, Washington, D. C.
- Tsurutani, B. T., W. D. Gonzalez, A. L. C. Gonzalez, F. Tang, J. K. Arballo, and M. Okada (1995), Interplanetary origin of geomagnetic activity in the declining phase of the solar cycle, *J. Geophys. Res.*, 100, 21,717–21,733.
- Wang, Y.-M., and N. R. Sheeley Jr. (1997), The high-latitude solar wind near sunspot maximum, *Geophys. Res. Lett.*, 24, 3141–3144.
- Zhao, X.-P., and J. T. Hoeksema (1997), Is the geoeffectiveness of the January 6, 1997, CME predictable from solar observations?, *Geophys. Res. Lett.*, 24, 2965–2968.
- Zhao, X.-P., and J. T. Hoeksema (1998), Central axial field direction in magnetic clouds and its relation to southward interplanetary magnetic field events and dependence on disappearing solar filaments, *J. Geophys. Res.*, 103, 2077–2083.

R. L. McPherron, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095-1567, USA. (rmcpherron@igpp.ucla.edu)

G. Siscoe, Center for Space Physics, Boston University, Boston, MA 02215, USA. (siscoe@bu.edu)